IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION PAPERS

OF

CHRISTOPHER ANDREW BARTON

FOR

UPDATING COMPUTER FILES

BACKGROUND OF THE INVENTION

Field of the Invention

This invention relates to the field of computers. More particularly, this invention relates to mechanisms for the updating of computer files.

Description of the Prior Art

A problem in the field of computing is the requirement for regular updating of computer files (possibly by downloading a complete new version of the file or by an incremental update in which modifying data for modifying the existing file to form the updated version is downloaded) held on many different computers. A software update may be required because the program has altered in response to the occurrence of bugs within the program or to add additional functionality to a program. Another need for frequent computer file updates is when the computer file represents rapidly evolving data needed by the computer. An example of this is the computer virus definitions data that is used by many anti-virus computer programs. This computer virus definition data is typically updated when a new virus is encountered such that the anti-virus software may provide counter-measures to the new virus. In order that the anti-virus software being used may operate in an effective manner it is important that it should use the most up to date virus definition data.

In response to this need, anti-virus software suppliers often provide download facilities from which users can download the most up to date versions of the computer virus definition data. One problem with this approach is that a user must know that an updated virus definition data file is present in order that it should be downloaded. One way to deal with this is to configure the computer program software to automatically check for new computer virus definition data at periodic intervals. If these intervals are made too short, then this presents an unnecessary burden upon the computer systems involved. Conversely, if the intervals are made too large, then a significant update required to deal with a new virus threat may not be downloaded in sufficient time to adequately protect from that virus threat.

A further problem associated with the downloading of virus definition data from the anti-virus software supplier is that peak demand for the download of the new data may cause the systems to malfunction. Computer viruses are becoming increasingly common and destructive. With this background, the release of a new computer virus attracts considerable media attention resulting in many users

simultaneously trying to download the updated data in a manner that causes this process to fail.

Various techniques for updating software are disclosed in US-A-5,940,074, US-A-4,763,271, US-A-5,919,247, US-A-5,577,244, US-A-5,809,287, US-A-5,933,647 and US-A-5,732,275. A technique for updating anti-virus DAT files via a "push" method is disclosed in US-A-6,035,423.

SUMMARY OF THE INVENTION

Viewed from one aspect the present invention provides a computer program product for updating a computer file used by a computer, said computer program product comprising:

(i) tag detecting code operable to detect within data received by said computer a tag indicative of existence of an updated version of said computer file; and

(ii) update triggering code operable upon detection of said tag to trigger downloading from a predetermined source to provide said updated version of said computer file for use by said computer.

The invention uses received data (such as e-mail messages) to provide peer-to-peer notification of the existence of an updated version of a computer file. In this way, the existence of an updated version of a computer file may be automatically disseminated in a manner that provides a degree of smoothing in the level of demand on the downloading systems for updating the computer file. Furthermore, those computers that receive more data from elsewhere are more likely to receive notification of an updated version of a computer file sooner than those computers receiving little data from elsewhere. Computers receiving much data from elsewhere are typically highly active and accordingly are the ones where early receipt of the updated computer file will be most beneficial. As an example, in the context of the distribution of computer virus definition data, computers receiving many e-mail messages are often the ones at high risk of computer virus infection and so the technique of the invention that results in these computers being likely to be among the first to be updated matches well with a priority based upon the risk from a new computer virus.

The invention recognises that with the widespread adoption of data transfer mechanisms, such as e-mail messaging, between computers, these mechanisms can be effectively used to disseminate information regarding computer file updates with very little additional overhead. In particular, in preferred embodiments the tag may be

embedded within a header portion of an e-mail message. Such header portions already include other information typically giving the transmission history of the message, which is not normally acted upon by the user and thus simplifying the embedding within this header of notification of computer file updates. This embedding may be done, for example, by any of a mail gateway, a server or workstation.

A particularly effective form of tag is one which indicates the version level of a particular computer file as this lends itself to ready comparison with other data and matches the way in which computer file suppliers already characterise their updates.

A computer receiving an e-mail message may also inserts its own tag into the message if it is itself using a computer file more up-to-date than any tag already present within the e-mail message. The computer could replace any existing tag already within the e-mail message with its indication of the more up to date version of that computer file, or alternatively it may simply add an additional tag indicating the existence of that more up to date version of the computer file. In either case, this feature significantly enhances the way in which the notification of the existence of an updated computer file will cascade through a network of many linked computers via e-mail messages, i.e. once a particular computer has itself been updated, then it may serve to add the notification of the existence of that update to any further e-mail messages that pass through it.

Whilst the present invention is suitable for the distribution of many different types of computer files, it is particularly well suited to use in anti-virus systems whereby the computer file may be a computer virus definition file, a computer virus detection engine file or an anti-virus computer program file. In each of these cases, the automatic notification of the existence of an update in a manner that can trigger the download of that update whilst avoiding excessive peaks in download demand is strongly advantageous.

The predetermined source from which a computer can download its updated version may take many forms. Within a local area network, a particular client computer may be configured to seek its updates from a specified location on a local server computer. A local server computer may be configured to seek its updates via the internet from a specified FTP source. There are numerous other possibilities.

In order to reduce the likelihood of malicious use of this updating mechanism, preferred embodiments are one in which the tag data is encrypted to make its misuse

more difficult. A further safeguard that might be used instead or as well as encryption that an update will not be triggered if the difference in the version level indicated is too large (e.g. greater than 1000 version levels) is as this may well be indicative of malicious intervention or malfunction.

Preferred embodiments also seek to smooth peak download demand by introducing an initial update delay period following detection of the tag. This update delay period can be varied such that a server version of the program will attempt to download very rapidly as the risk for a server is large whereas a workstation version may have a longer initial delay before attempting a download. The delay can be configured for each computer to match the risk associated with that computer using an out-of-date file.

Should the download fail, then the problem of an excessive build-up in requests for download may be reduced by the introduction of a failure delay period before the computer retries to download the new version of the computer file. This failure delay period may be made psuedo-random in order to try to reduce any pathological condition with competing computers simultaneously reissuing their requests for download.

Viewed from another aspect the present invention provides a computer program product comprising:

(i) tag inserting code operable to insert a tag within data indicating a version level of a computer file used by a first computer, said tag being operable to trigger an update of an older version of said computer file used by a second computer when said data is received by said second computer.

It will be appreciated that some elements within a computer network may not themselves use a particular computer file and yet may be usefully integrated within the scheme of this invention by being configured to insert suitable tags within E-mail messages passing through them.

The present invention also provides a method of updating a computer file and an apparatus for updating a computer file in accordance with the appended claims.

The above, and other objects, features and advantages of this invention will be apparent from the following detailed description of illustrative embodiments which is to be read in connection with the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1, 2 and 3 illustrate operation of the technique of the present invention in distributing notification of the existence of an update and triggering download of an update;

Figure 4 is a flow diagram showing the processing performed by a peer computer of the type illustrated in Figures 1 to 3;

Figure 5 schematically illustrates a tag of the type that may be inserted within an e-mail message header; and

Figure 6 illustrates a computer apparatus that may be used to implement the technique of the present invention.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

Figure 1 shows a plurality of computers linked via the internet 2 (or some other communication link). A software supplier provides an FTP server 4 from which updated versions of a computer file may be downloaded. Figure 1 illustrates a first local area network 6 and a second local area network 8 linked to the internet 2. The first local area network 6 includes a mail gateway computer 10, a local server 12 and two client workstations 14, 16. In a similar way the second local area network 8 includes a mail gateway 18, a local server 20 and two client workstations 22, 24.

Using known e-mail protocols and computer programs, the various workstation computers 14, 16, 22, 24 shown in Figure 1 exchange e-mail messages. In alternative embodiments the data exchange could take other forms, such as internet web pages or word processing files that contain headers or the like in which tags for triggering updates may be embedded. All of the computers illustrated in Figure 1 utilise anti-virus software that requires access to an up to date computer virus definition data file. As users of this computer virus definition data file the various mail gateway, server and workstation computers are peers. The illustrated starting situation in Figure 1 is that all of the computers are running the most up to date version (#3) of the computer file that is the version currently stored on the FTP server 4.

Figure 2 illustrates the start of the dissemination of an update. The software supplier loads an updated version of the computer virus definition data file onto the FTP server 4. A manual or an automatic timed update of the mail gateway 10 then takes place that updates the version of the computer file on the mail gateway to #4. Subsequently, an e-mail message is issued by the workstation computer 16 destined for the workstation computer 24. The anti-virus software within the workstation

computer 16 tags the e-mail header of the e-mail message with the version #3 of the computer virus definition data file that the workstation computer 16 is currently using. This e-mail message first passes through the local server 12 which is also using this same version #3 and so leaves the tag unaltered. When the e-mail message reaches the mail gateway 10, the mail gateway checks the tag within the e-mail message header and determines that it is itself using a more up-to-date version of the computer file in question and so replaces the tag with one that indicates that version #4 has been used and is available. The tag may also indicate whether or not that computer may itself serve as the source of the update data for other computers. This e-mail message then propagates via the internet 2 and through the mail gateway 18 and the local server 20 until it reaches the workstation computer 24. Each of the mail gateway 18, the local server 20 and the workstation computer 24 examines the tag within the e-mail message header and determines that it indicates the existence of a more up to date version of the computer virus definition data file than that which they are currently using. Accordingly, each of the mail gateway 18, the local server 20 and the workstation computer 24 is triggered (after an initial delay period) to download the updated version of the computer file from the FTP server 4 (a predetermined source of updates). The "*" indicates that a particular peer computer has detected that it should download an updated version of the computer file. The mail gateway 18, the local server 20 and the workstation computer 24 will continue to use version #3 until they have the updated version #4. The update mechanism used is preferably the pre-existing standard "pull" mechanism, but it is envisaged that alternative emergency or special purpose update mechanisms for use when triggered by a received tag could be provided.

Figure 3 shows the situation when the updates to the mail gateway 18, the local server 20 and the workstation computer 24 have all taken place. In the example shown, the user of the workstation computer 24 sends an e-mail message to the user of the workstation computer 16 that is also copied to the user of the workstation computer 22. As the workstation computer 24 that is the source of this e-mail message has now been updated to version #4, it includes within the header of the e-mail message a tag indicating that version #4 exists. When the workstation computer 22 receives this message, this is detected as indicating that the workstation computer 22 should download the updated version #4 from the FTP server 4. The e-mail message also propagates via the local server 20, the mail gateway 18 and the mail

gateway 10 towards the other target recipient that is the workstation computer 16. All of these peer computers have already been updated, thus they do not action version #4 tags and below. The first computer reached in this transmission path that has not yet been updated is the local server 12 and the second is the workstation computer 16. Both of these computers also detect that the tag shows a version level #4 higher than that which they are currently using #3 and accordingly trigger the download of an update from the FTP server 4.

It will be appreciated that the mechanisms shown in Figures 1, 2 and 3 allow for the automatic and rapid dissemination of the information that an updated file exists. Furthermore, those computers that receive more e-mail messages are more likely to receive this notification sooner. These are the very computers that are generally at a high risk from computer viruses and accordingly it is appropriate that they should be the first that download the updated version of the computer virus definition data file. A little used computer will only download its update at a later time, and yet this will pose potentially lower level of risk since the little used computer is less likely to receive an infectious element.

Figure 4 is a flow diagram schematically illustrating the processing performed by a particular peer computer.

At step 26, the computer receives an e-mail message. At step 28 the computer searches the message header of the e-mail for any header tag present and decrypts this if it is found. In some embodiments the header tag may not be encrypted.

At step 30, a test is made as to whether any header tag has been found. If a header tag has not been found, then the e-mail message is scanned for computer viruses using the existing anti-virus software at its current level of update and a header tag added to the e-mail message indicative of that current level of update at step 32.

If a header tag is found at step 30, then step 34 tests whether or not that header tag indicates a version of the software that is older than that held by the computer performing the process illustrated in Figure 4. If the received header tag is indicative of an older version, the processing proceeds to step 32 at which the e-mail message is scanned using what is known to be more up to date data than has previously been applied to that message and the header tag indicative of that more up to date data is added to the e-mail header. The existing tag indicative of the older version of the computer file may or may not be removed. The tag may also include parameters

indicative of previous processing applied to the message, such as the program options set (e.g. all files, macro heuristics, all heuristics) on previous anti-virus scans applied to the message. These parameters can be used to determine whether or not further scanning is to be applied by the computer currently processing the message.

If the test at step 34 indicates that the received e-mail message included a header tag that was not older than the local version, then processing proceeds to step 36. Step 36 tests whether or not the tag of the received e-mail message indicates a newer version of the computer file is available. If a newer version is not available, then processing proceeds to step 37. Step 37 decided whether or not the message should be scanned at step 32 in dependence upon parameters set on the processing computer and parameters within the tag as mentioned above that indicate in more detail what previous scanning has been applied to the message. If the test at step 36 indicates that a newer version of the computer file than that stored by the local computer is available, then processing proceeds to step 40.

Step 40 tests how may versions ahead of the current version the tag within the received e-mail message indicates is available. If this number exceeds a predetermined threshold N, then this is indicative of some malfunction or malicious interference with the message tags and accordingly processing proceeds to step 32.

If the test at step 40 indicates that the updated version is less than the threshold number N ahead of the currently used version, then processing proceeds to step 42 which scans using the currently held version and then imposes an initial delay before processing proceeds to step 44 where an update attempt is triggered to download the updated version of the computer file from a remote source. The remote source may be an FTP server 4 linked via the internet 2 (or any other link) to the computer in question, or could be a server file location within a local area network 6, 8 or some other source.

At step 46, a test is made as to whether or not the update attempt has failed. If the update attempt has not failed, then processing continues with other normal e-mail processing operations at step 38.

If the update attempt has failed, then processing proceeds to step 48 which imposes a psuedo-random failure delay prior to returning processing to step 44 to attempt another update. A predetermined number of update attempt failures may trigger a user warning message.

Figure 5 illustrates a message tag of the type that may be inserted within an E-mail message header. The "X-" prefix in the tag is one defined some in standard e-mail protocols as indicating that the information that follows on the line is for information purposes only and is not actively processed in normal systems. In the illustrated example, the tag starts with a coding for "McAfee-Sig:". This is a code sequence that is searched for by peer computers within e-mail message headers to detect the presence of a tag indicating what version levels of an anti-virus system have already been applied to that e-mail message. This version information follows in the form of "<S#-xxxE#-yyyD#-zzz>", portions of which respectively indicate the anti-virus software program version number, the computer virus detection engine version number and the computer virus definition data version number. In practice, this version information may be encrypted to make it more difficult to tamper with the information in an attempt to misdirect or interfere with the update mechanisms. Various known encryption techniques may be used. The tag could also include parameters indicating previously applied scan options or other data and could extend over more than one line.

Figure 6 schematically illustrates a computer 50 of the type that may be used to implement the techniques described above (more simple appliance type devices could also be used). The computer 50 includes a central processing unit 52, a read only memory 54, a random access memory 56, a network link 58, a hard disk drive 60, a display driver 62, a display 64, a user input/output driver 64, a keyboard 66 and a mouse 68.

In operation, the central processing unit 52 executes computer programs stored upon the hard disk drive 60 or within the read only memory 54 using the random access memory as working memory. User inputs for controlling the computer 50 are received from the keyboard 66 and the mouse 68 via the user input/output unit 64. Processing results may be displayed to the user using the display 64 via the display driver unit 62.

In operation, an e-mail message may be received from the internet 2 via the network link 58 into an e-mail program being executed by the central processing unit 52. Part of this e-mail program may trigger an anti-virus scan of the received e-mail. This anti-virus scan includes the processing illustrated in Figure 4 and accordingly detects if the received e-mail message includes the information that an updated version of any of the components of the anti-virus system exists for download.

Should such updated versions exist, then they may be downloaded by the computer 50 under program control from a remote source, such as an FTP server 4, via the internet 2 and the network link 58.

The updated computer files, such as the anti-virus software program, the search engine program or the virus definitions, will typically then be stored on the hard disk drive 60 of the computer 50. The program that causes the processing of Figure 4 to take place will also typically be stored upon the hard disk drive 60. The computer program may be distributed via a recordable medium, such as a floppy disk or a CD, or may itself be downloaded via the network link 58.

The computer 50 may pass the e-mail message onto another computer or may itself originate a new e-mail message. In either case, any outbound e-mail message is marked with a tag indicating the version levels of the anti-virus software components used by the computer 50 if these are more up to date than any indications already within the e-mail message.

Although illustrative embodiments of the invention have been described in detail herein with reference to the accompanying drawings, it is to be understood that the invention is not limited to those precise embodiments, and that various changes and modifications can be effected therein by one skilled in the art without departing from the scope and spirit of the invention as defined by the appended claims.